A new method for constructing small-bias spaces from Hermitian codes

Olav Geil[1], Stefano Martin[1], and Ryutaroh Matsumoto[1,2]

[1] Department of Mathematical Sciences, Aalborg University, Denmark
olav@math.aau.dk,
stefano@math.aau.dk,
[2] Department of Communications and Integrated Systems, Tokyo Institute of Technology, Japan
ryutaroh@rmatsumoto.org

Abstract. We propose a new method for constructing small-bias spaces through a combination of Hermitian codes. For a class of parameters our multisets are much faster to construct than what can be achieved by use of the traditional algebraic geometric code construction. So, if speed is important, our construction is competitive with all other known constructions in that region. And if speed is not a matter of interest the small-bias spaces of the present paper still perform better than the ones related to norm-trace codes reported in [12].

Keywords. Small-bias space, balanced code, Gröbner basis, Hermitian code.

1 Introduction

Let $\boldsymbol{X} = (X_1, \ldots, X_k)$ be a random vector that takes on values in $\mathbb{F}_2^k$. As shown by Vazirani [17] the variables $X_1, \ldots, X_k$ are independent and uniformly distributed if and only if

$$\text{Prob}\left(\sum_{i \in T} X_i = 0\right) = \text{Prob}\left(\sum_{i \in T} X_i = 1\right) = \frac{1}{2} \tag{1}$$

holds for every non-empty set of indexes $T \subseteq \{1, \ldots, k\}$. In particular, if (1) is to hold for a space $\mathcal{X} \subseteq \mathbb{F}_2^k$ then necessarily $\mathcal{X}$ must be equal to $\mathbb{F}_2^k$. There is a need for much smaller spaces $\mathcal{X} \subseteq \mathbb{F}_2^k$ with statistical properties close to that of (1). In the following by a space we will mean a multiset $\mathcal{X}$ with elements from $\mathbb{F}_2^k$ (this we write $\mathcal{X} \subseteq \mathbb{F}_2^k$). The multiset $\mathcal{X}$ is made into a probability space by adjoining to each element $\boldsymbol{x} \in \mathcal{X}$ the probability $p(\boldsymbol{x}) = i(\boldsymbol{x})/|\mathcal{X}|$ where $i(\boldsymbol{x})$ denotes the number of times $\boldsymbol{x}$ appears in $\mathcal{X}$. As a measure for describing how close a given space $\mathcal{X}$ is to the above situation with respect to randomization, Naor and Naor [15], and Alon et. al. [1] introduced the concept of ϵ-biasness [15, Def. 3]. (See also [14]).

Definition 1. *A multiset $\mathcal{X} \subseteq \mathbb{F}_2^k$ is called an ϵ-bias space if*

$$\frac{1}{|\mathcal{X}|} \left| \sum_{\boldsymbol{x} \in \mathcal{X}} (-1)^{\sum_{i \in T} x_i} \right| \leq \epsilon \tag{2}$$

holds for every non-empty index set $T \subseteq \{1, \ldots, k\}$.

Clearly, the ϵ in Definition 1 can be taken to be a number between 0 and 1. Good randomization properties are achieved when ϵ is close to 0 as (2) becomes (1) when $\epsilon = 0$. Multisets with ϵ small are called small-bias spaces. They are useful as sample spaces in applications such as automated theorem proving, derandomization of algorithms, program verification, and testing of combinatorial circuits. Rather than saying that a multiset is an ϵ-bias space we will often just say that it is ϵ-biased. Another name for ϵ-bias space is ϵ-bias set [2, Def. 1] and [12, Def. 1.1]. This notion may be a little misleading as the item under consideration is actually a multiset. One way of constructing small-bias spaces is through the use of error-correcting codes.

Definition 2. *A binary $[n, k]$ code is said to be ϵ-balanced if every non-zero code word $\boldsymbol{c}$ satisfies*

$$\frac{1-\epsilon}{2} \leq \frac{w_H(\boldsymbol{c})}{n} \leq \frac{1+\epsilon}{2}.$$

Here $[n, k]$ means that the code is linear, of dimension k and length n. Further, w_H denotes the Hamming weight.

There is a simple direct translation [1] between the concepts described in Definition 1 and Definition 2:

Theorem 1. *Let G be a generator matrix for an ϵ-balanced binary $[n, k]$ code. The columns of G constitute an ϵ-bias space $\mathcal{X} \subseteq \mathbb{F}_2^k$ of size n. Similarly, using the elements of an ϵ-bias space $\mathcal{X}$ as columns of a generator matrix an ϵ-balanced code is derived.*

The following example illustrates the above theorem. It also shows why it is important in Definition 1 to work with multisets rather than sets.

Example 1. Consider the matrix

$$G = \begin{bmatrix} 0\,1\,0\,1\,0\,1\,0\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,0\,0\,0\,0 \end{bmatrix}.$$

The code having G as a generator matrix is ϵ-balanced with $\epsilon = 1/3$ and indeed the multiset made from the columns of G is $\epsilon = 1/3$ biased. Treating the columns as a set (rather than a multiset) we derive

$$\mathcal{X}' = \{(0,0,1), (1,0,1), (0,1,1), (1,1,1), (0,0,0)\}.$$

The smallest value of ϵ for which $\mathcal{X}'$ is ϵ-biased is $\epsilon = 3/5$.

A standard construction from [1] tells us how to make small-balanced codes (meaning ϵ-biased codes with ϵ small):

Theorem 2. *Let $q = 2^s$ for some integer s and consider a q-ary $[N, K, D]$ code C. Let C_s be the (binary) $[2^s, s]_2$ Walsh-Hadamard code, $s \geq 1$. The concatenated code derived by using C as outer code and C_s as inner code is an $\epsilon = (N-D)/N$-balanced binary code of length $n = N2^s$ and dimension $k = Ks$.*

Proof. The result relies on the fact that every non-zero codeword of C_s contains exactly as many 0s as 1s.

The literature contains various examples of small-bias spaces that cannot all be compared to each other. We refer to [2, Sec. 1] for more details. In the following we will concentrate on important families of multisets for which comparison can be made. We remind the reader of how bigO notation works when given functions of multiple variables. In our situation we have real valued positive functions $f_i(x, y), i = 1, 2$ where x can take on any value in $\mathbb{Z}^+$ but for every fixed choice of x the variable y can only take on values in an interval $I(x) \subseteq \mathbb{R}^+$. By $f_1(x, y) = \mathcal{O}(f_2(x, y))$ we mean that a witness (C, κ) exists such that for all x with $\kappa < x$ and all $y \in I(x)$ it holds that $f_1(x, y) \leq C f_2(x, y)$. We are interested in upper bounding the size of $\mathcal{X}$ which will be done in terms of bigO estimates as above. At the same time we are interested in lower bounding the length of the words in the multiset $\mathcal{X}$. Such estimates are described using bigOmega notation. We remind the reader that by definition $f(x) = \Omega(g(x))$ if and only if $g(x) = \mathcal{O}(f(x))$. As we are only interested in bigOmega estimates the meaning of k changes accordingly. We have the following results:

- Using Reed-Solomon codes as outer codes in Theorem 2 one achieves [1,2] for all possible choices of ϵ and k

$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\frac{k^2}{\epsilon^2 \log^2(k/\epsilon)}\right).$$

  This is called the RS-bound.

- Let $P_1, \dots, P_{\mathcal{N}-1}, Q$ be rational places of an algebraic function field over $\mathbb{F}_q$ and denote by g the genus. Assume $\mathcal{N} = (\sqrt{q}-1)g$. That is, we assume that the function field attains the Drinfeld-Vladut bound. Using codes $C_{\mathcal{L}}(U = P_1 + \cdots + P_{\mathcal{N}-1}, mQ)$ with $g < m$ as outer codes one gets for all ϵ and k (see Section 2 for a discussion)
$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\frac{k}{\epsilon^3 \log(1/\epsilon)}\right).$$
This result which is in the folklore is known as the AG-bound.
- Using Hermitian codes with $m < g$ as outer codes one achieves [2] for $\epsilon \geq k^{-\frac{1}{2}}$
$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\left(\frac{k}{\epsilon^2 \log(1/\epsilon)}\right)^{\frac{5}{4}}\right). \tag{3}$$
This we call the BT-bound after the authors of [2], Ben-Aroya and Ta-Shma.
- Using in larger generality Norm-Trace codes of low dimension as outer codes one achieves [12] for $l = 4, 5, \dots$ and $\epsilon \geq k^{-\frac{1}{\sqrt{l}}}$ (see Section 5)
$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\left(\frac{k}{\epsilon^{l-\sqrt{l}} \log(1/\epsilon)}\right)^{\frac{l+1}{l}}\right).$$
Here, $l = 4$ corresponds to the Hermitian case described in [2].
- The Gilbert-Varshamov bound also applies to the small-bias spaces (as usual in a non-constructive way). It is derived by plugging into the Gilbert-Varshamov bound for binary codes $d = n/2$ and to make a Taylor approximation on the resulting formula. The construction uses Theorem 1 directly. It guarantees for all ϵ and k the existence of multisets with
$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\frac{k}{\epsilon^2}\right).$$
- The linear programming bound tells us that we cannot hope to produce ϵ-bias spaces with
$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \quad |\mathcal{X}| = \mathcal{O}\left(\frac{k}{\epsilon^2 \log(1/\epsilon)}\right).$$

One way of comparing the above results is to choose $\epsilon = k^{-\alpha}$, $\alpha \in \mathbb{R}^+$ and then to take the logarithm with base k. The bigO notation suggests that we then let k go to infinity. The origin of this point of view is [2, Sec. 1].

When making the above operation we must be careful to specify which choices of α are allowed. We remind the reader of the little-o notation. Given functions $f_i(x) : \mathbb{Z}^+ \to \mathbb{R}^+, i = 1, 2$ by $f_1(x) = o(f_2(x))$ we mean that for every choice of $c \in \mathbb{R}^+$ there exists a $\kappa(c) \in \mathbb{Z}^+$ such that when $\kappa(c) < x$ then necessarily $f_1(x) \leq cf_2(x)$. We have:

- RS-bound: The family of concatenated codes from Theorem 2 with Reed-Solomon codes as outer codes gives
$$\log_k(|\mathcal{X}|) = 2 + 2\alpha + o(1)$$
for all choices of $\alpha \in \mathbb{R}^+$.
- AG-bound: The family of concatenated codes from Theorem 2 with algebraic geometric codes as outer codes and $g < m$ gives
$$\log_k(|\mathcal{X}|) = 1 + 3\alpha + o(1)$$
for all choices of $\alpha \in \mathbb{R}^+$.
- BT-bound: The family of concatenated codes from Theorem 2 with Hermitian codes as outer codes and $m < g$ gives
$$\log_k(|\mathcal{X}|) = \frac{5}{4} + \frac{5}{2}\alpha + o(1)$$
for all choices of $\alpha \in]1/2, \infty[$.
- The family of concatenated codes from norm-trace codes of low dimension gives
$$\log_k(|\mathcal{X}|) = \frac{l+1}{l}(1 + \alpha(l - \sqrt{l})) + o(1)$$
for $l = 4, 5, \ldots$, and for all $\alpha \in [1/\sqrt{l}, \infty[$ (see Section 5).
- The Gilbert-Varshamov bound and the Linear Programming bound in combination tell us that we can achieve
$$\log_k(|\mathcal{X}|) = 1 + 2\alpha + o(1)$$
for all choices of $\alpha \in \mathbb{R}^+$ but no better than this.

In the present paper we shall introduce a new family of small-bias spaces using a combination of Hermitian codes as outer code. This family gives

$$\log_k(|\mathcal{X}|) = \frac{4}{3} + \frac{8}{3}\alpha + o(1)$$

for all choices of $\alpha \in \mathbb{R}^+$. We allow $2g < m$ and it is therefore surprising that for $\alpha \in]1, \infty[$ the achievements are better than those of the Hermitian codes with $g < m$. Our small-bias spaces perform better than the ones derived from norm-trace codes for all $l \geq 5$ (see Section 5 for the proof). For $\alpha < 1$ they behave better than what can be achieved using Reed-Solomon codes as outer code. For $\alpha < 1$ admittedly the new ϵ-bias spaces perform worse than the spaces coming from the AG construction. This, however, is only part of the picture. It turns out that to construct the spaces with $\alpha < 1/2$ from the AG construction requires quite a number of operations. In contrast, our construction is considerable faster. We shall revert to this issue in Section 4. Before dealing with the new construction we will investigate how to ensure $\epsilon = k^{-\alpha}$ in the case of the AG bound. It turns out that for $\alpha < 1/2$ the situation is rather complicated. We include the description here, as to our best knowledge, the details cannot be found in the literature.

2 The AG-bound

Let q be a power of 2 and consider an algebraic function field over $\mathbb{F}_{q^2}$ of genus g with at least $\mathcal{N} = (q-1)g$ rational places. That is, the function field attains the Drinfeld-Vladut bound. As noted in the introduction Theorem 2 equipped with a one-point algebraic geometric code from the above function field produces ϵ-bias spaces $\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}$ with

$$|\mathcal{X}| = \mathcal{O}\left(\frac{k}{\epsilon^3 \log_2(\frac{1}{\epsilon})}\right). \tag{4}$$

In the following we investigate how to achieve corresponding values ϵ and k under the requirement $\epsilon = k^{-\alpha}$, $\alpha > 0$, and $k \to \infty$. Observe, that in this situation for any fixed α we have $\epsilon \to 0$. For completeness we start by proving (4) in this setting.

Consider rational places $P_1, \ldots, P_{\mathcal{N}-1}, Q$ and let $U = P_1 + \cdots + P_{\mathcal{N}-1}$ and $G = (ag)Q$ with $a \geq 1$. The code $C_{\mathcal{L}}(U, G)$ has parameters $N = (q-1)g-1$, $K \geq \deg G - g = (a-1)g$, and $D \geq N - \deg G = ((q-1)-a)g-1$. As we are interested in asymptotics we shall assume $N = (q-1)g$ and $D \geq ((q-1)-a)g$. From Theorem 2 we get ϵ-bias spaces with $\epsilon = a/(q-1)$, $\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}$. Here, $k = 2\log_2(q)(a-1)g$ and we have $|\mathcal{X}| = q^2 N = (q^3 - q^2)g$. As a is bounded below by 1 and $\epsilon \to 0$ we need $q \to \infty$ when $k \to \infty$. So the task basically boils down to establishing a sequence of function fields over increasingly large fields and a corresponding function

$a(q)$ such that

$$|\mathcal{X}| = \mathcal{O}\left(\frac{2\log_2(q)(a-1)g}{\left(\frac{a}{q-1}\right)^3 \log_2\left(\frac{q-1}{a}\right)}\right). \tag{5}$$

Note that the argument on the right side is a function in the single variable q as by construction now g is a function of q. We have

$$\frac{2\log_2(q)(a-1)g}{\left(\frac{a}{q-1}\right)^3 \log_2\left(\frac{q-1}{a}\right)} \geq \frac{1}{2}\frac{\log_2(q)(a-1)}{a^3(\log_2(q-1)-\log_2(a))}|\mathcal{X}|$$

as $(q-1)^3 \geq \frac{1}{4}(q^3-q^2)$ holds for $q \geq 2$. In conclusion (5) holds if $a(q) = \mathcal{O}(1)$.

We first assume that the sequence of function fields are the Hermitians which are function fields with $g = q(q-1)/2$. Here, actually the number of rational places is $2qg + q^2 + 1$ but we shall only use $(q-1)g$ of them. Let $a = 1 + q^{-c}$ where $0 \leq c < 2$. Clearly, $a(q) = \mathcal{O}(1)$ as requested. We have $k = 2\log_2(q)q^{-c}g = q^{2-c}q^{\beta}$ where $\beta(q) \to 0$ for $q \to \infty$. Hence, asymptotically $\epsilon = k^{-\alpha}$ with $\alpha = 1/(2-c)$. In other words the situation is clear for $\alpha \in [\frac{1}{2}, \infty[$.

To achieve $\alpha \in]0, \frac{1}{2}[$ is more difficult. The problem is to keep $a(q) = \mathcal{O}(1)$ at the same time as having $\epsilon = k^{-\alpha}$. For this purpose we consider families of towers of function fields over $\mathbb{F}_{q^2}$ attaining the Drinfeld-Vladut bound [5]. We will need one tower for each value of q. Note that in such a tower for arbitrary $v \geq 2$ we can find a function field with $g \geq q^v$. Say $g = q^{v+d(q)}$, where $d(q) \geq 0$ holds. Let $a(q) = 1 + q^{-d(q)}$ then clearly $a(q) = \mathcal{O}(1)$ holds. We have $k = 2\log_2(q)(a-1)g = q^{v+\beta}$ where $\beta(q) \to 0$ for $q \to \infty$. Also $\epsilon = q^{-1+\gamma}$ where $\gamma(q) \to 0$ for $q \to \infty$. Hence, $k^{-\alpha} = \epsilon$ asymptotically means $v\alpha = 1 \Rightarrow \alpha = 1/v$. As we only assumed $v \geq 2$ we have established that all $\alpha \in]0, \frac{1}{2}[$ can be attained.

For our purpose the best candidate for a family of good towers of function fields is the second construction by Garcia and Stichtenoth [5]. In [16] it was shown how to construct $C_{\mathcal{L}}(U, G)$ codes from this tower using

$$\mathcal{O}\left((N\log_q(N))^3\right) \tag{6}$$

operations over $\mathbb{F}_{q^2}$. Although we might only need codes of small dimension the method as stated requests us to find bases for all one-point codes. As shall be demonstrated in Section 4 the small-bias spaces of the present paper can be constructed much faster than what (6) guarantees for the AG construction.

3 The new small-bias spaces

In the present paper we propose a new choice of outer codes in the construction of Theorem 2. As already mentioned this results in small-bias spaces with good properties. The new choice of outer codes is derived by combining two Hermitian codes as described below. The easiest way to explain the combination is by using the language of affine variety codes [4] and we therefore start our investigations with a presentation of Hermitian codes as such.

Definition 3. *Given a monomial ordering $\prec$ and an ideal $I \subseteq \mathbb{F}[X_1, \ldots, X_m]$ (here $\mathbb{F}$ is any field) the footprint is*

$$\Delta_{\prec}(I) := \{X_1^{\alpha_1} \cdots X_m^{\alpha_m} \mid X_1^{\alpha_1} \cdots X_m^{\alpha_m} \ \textit{is not a leading monomial} \\ \textit{of any polynomial in} \ \ I\}.$$

We have the following two useful results [3, Pro. 4 and Pro. 8, Sec. 5.3].

Theorem 3. *The set $\{M+I \mid M \in \Delta_{\prec}(I)\}$ is a basis for $\mathbb{F}[X_1, \ldots, X_m]/I$ as a vector space over $\mathbb{F}$.*

As a corollary one gets the following result often referred to as the footprint bound [7,9].

Theorem 4. *Assume I is zero-dimensional (meaning that $\Delta_{\prec}(I)$ is finite). The variety $\mathbb{V}_{\bar{\mathbb{F}}}(I)$ satisfies $|\mathbb{V}_{\bar{\mathbb{F}}}(I)| \leq |\Delta_{\prec}(I)|$.*

Consider the Hermitian polynomial $X^{q+1} - Y^q - Y$ and the corresponding ideal

$$I = \langle X^{q+1} - Y^q - Y \rangle \subseteq \mathbb{F}_{q^2}[X, Y].$$

Define a monomial function w by $w(X) = q$ and $w(Y) = (q+1)$ and consider the weighted degree monomial ordering $\prec_w$ given by $X^{\alpha_1}Y^{\beta_1} \prec_w X^{\alpha_2}Y^{\beta_2}$ if one of the following two conditions holds:

1. $w(X^{\alpha_1}Y^{\beta_1}) < w(X^{\alpha_2}Y^{\beta_2})$.
2. $w(X^{\alpha_1}Y^{\beta_1}) = w(X^{\alpha_2}Y^{\beta_2})$ but $\beta_1 < \beta_2$.

Observe for later use that no two different monomials in

$$\Delta_{\prec_w}(I) = \{X^i Y^j \mid 0 \leq i \ \text{ and } \ 0 \leq j < q\}$$

are of the same weight implying that $w : \Delta_{\prec_w}(I) \rightarrow \langle q, q+1 \rangle$ is a bijection. Observe also that the Hermitian polynomial $X^{q+1} - Y^q - Y$ contains exactly two monomials of highest weight. The implication of this is that

$$w(\text{lm}(F(X, Y)) = w(\text{lm}(F(X, Y) \text{ rem } \{X^{q+1} - Y^q - Y\})$$

holds for any polynomial $F(X,Y)$ that possesses exactly one monomial of highest weight in its support.
Consider next the ideal

$$I_{q^2} := \langle X^{q^2} - X, Y^{q^2} - Y \rangle + I.$$

The variety $\mathbb{V}_{\mathbb{F}_{q^2}}(I) = \mathbb{V}_{\mathbb{F}_{q^2}}(I_{q^2})$ consists of $n = q^3$ different points $\{P_1, \ldots P_n\}$. The set $\{X^{q^2} - X, X^{q+1} - Y^q - Y\}$ constitutes a Gröbner basis for I_{q^2} with respect to $\prec_w$ and therefore

$$\Delta_{\prec_w}(I_{q^2}) = \{X^i Y^j \mid 0 \leq i < q^2, 0 \leq j < q\}$$

holds. It now follows from Theorem 3 that

$$\{X^i Y^j + I_{q^2} \mid 0 \leq i < q^2, 0 \leq j < q\}$$

is a basis for $\mathbb{F}_{q^2}[X,Y]/I_{q^2}$ as a vector space over $\mathbb{F}_{q^2}$. The code construction relies on the bijective evaluation map $\mathrm{ev} : \mathbb{F}_{q^2}[X,Y]/I_{q^2} \rightarrow \mathbb{F}_{q^2}^n$ given by $\mathrm{ev}(F(X,Y) + I_{q^2}) = (F(P_1), \ldots, F(P_n))$. Theorem 4 tells us that we can estimate the Hamming weight of a word $\boldsymbol{c} = \mathrm{ev}(F(X,Y) + I_{q^2})$ by

$$w_H(\boldsymbol{c}) \geq n - |\Delta_{\prec_w}(\langle F(X,Y) \rangle + I_{q^2})|.$$

Without loss of generality we can assume $\mathrm{Supp}(F) \subseteq \Delta_{\prec_w}(I_{q^2})$. From the discussion prior to the definition of I_{q^2} we conclude that no two different monomials in $F(X,Y)$ are of the same weight. As a consequence

$$w(\mathrm{lm}(X^\alpha Y^\beta F(X,Y)) = w(\mathrm{lm}(X^\alpha Y^\beta F(X,Y) \text{ rem } \{X^{q+1} - Y^q - Y\})$$

holds for all $X^\alpha Y^\beta$. Write $\Lambda = w(\Delta_{\prec_w}(I)) = \langle q, q+1 \rangle$, $\Lambda^* = w(\Delta_{\prec_w}(I_{q^2})) \subseteq \Lambda$ and $\lambda = w(\mathrm{lm}(F)) \in \Lambda^*$. We have

$$|\Delta_{\prec_w}(\langle F(X,Y) \rangle + I_{q^2})| \leq |(\Lambda^* - (\lambda + \Lambda))| \leq |(\Lambda \backslash (\lambda + \Lambda)| = \lambda,$$

where the last equality comes from [10, Lem. 5.15]. Hence, $w_H(\boldsymbol{c}) \geq n - \lambda$ holds. Observe that

$$\Lambda^* = \{\lambda_1, \ldots, \lambda_g\} \cup \{2g, \ldots, n-1\} \cup \{\lambda_{n-g+1}, \ldots, \lambda_n\}, \tag{7}$$

where $\lambda_i \leq g - 1 + i$ for $i = 1, \ldots, g$. This is a general result for Weierstrass semigroups and not particular for the Hermitian function field. Having described the Hermitian codes as affine variety codes we are now ready to introduce the combination of codes on which our construction of small-bias spaces rely. Consider the ideal

$$I_{q^2}^{(2)} := \langle X_1^{q+1} - Y_1^q - Y_1, X_2^{q+1} - Y_2^q - Y_2, X_1^{q^2} - X_1, Y_1^{q^2} - Y_1, X_2^{q^2} - X_2, Y_2^{q^2} - Y_2 \rangle$$

and the corresponding variety

$$\mathbb{V}_{\mathbb{F}_{q^2}}(I_{q^2}^{(2)}) = \mathbb{V}_{\mathbb{F}_{q^2}}(I_{q^2}) \times \mathbb{V}_{\mathbb{F}_{q^2}}(I_{q^2}) = \{Q_1, \ldots, Q_{q^6}\}.$$

Define a monomial function $w^{(2)}$ given by $w^{(2)}(X_1) = (q, 0)$, $w^{(2)}(Y_1) = (q+1, 0)$,$w^{(2)}(X_2) = (0, q)$, and finally $w^{(2)}(Y_2) = (0, q+1)$. Let $\prec_{\mathbb{N}_0^2}$ be any monomial ordering on $\mathbb{N}_0^2$ and define $\prec_{w^{(2)}}$ by

$$X_1^{\alpha_1^{(1)}} Y_1^{\beta_1^{(1)}} X_2^{\alpha_1^{(2)}} Y_2^{\beta_1^{(2)}} \prec_w^{(2)} X_1^{\alpha_2^{(1)}} Y_1^{\beta_2^{(1)}} X_2^{\alpha_2^{(2)}} Y_2^{\beta_2^{(2)}}$$

if one of the following two conditions holds:

1. $w^{(2)}(X_1^{\alpha_1^{(1)}} Y_1^{\beta_1^{(1)}} X_2^{\alpha_1^{(2)}} Y_2^{\beta_1^{(2)}}) \prec_{\mathbb{N}_0^2} w^{(2)}(X_1^{\alpha_2^{(1)}} Y_1^{\beta_2^{(1)}} X_2^{\alpha_2^{(2)}} Y_2^{\beta_2^{(2)}})$
2. $w^{(2)}(X_1^{\alpha_1^{(1)}} Y_1^{\beta_1^{(1)}} X_2^{\alpha_1^{(2)}} Y_2^{\beta_1^{(2)}}) = w^{(2)}(X_1^{\alpha_2^{(1)}} Y_1^{\beta_2^{(1)}} X_2^{\alpha_2^{(2)}} Y_2^{\beta_2^{(2)}})$
   but
   $X_1^{\alpha_1^{(1)}} Y_1^{\beta_1^{(1)}} X_2^{\alpha_1^{(2)}} Y_2^{\beta_1^{(2)}} \prec_{\mathrm{lex}} X_1^{\alpha_2^{(1)}} Y_1^{\beta_2^{(1)}} X_2^{\alpha_2^{(2)}} Y_2^{\beta_2^{(2)}}$.

Here, $X_1 \succ_{\mathrm{lex}} Y_1 \succ_{\mathrm{lex}} X_2 \succ_{\mathrm{lex}} Y_2$ is assumed. The set $\{X_1^{q+1} - Y_1^q - Y_1, X_2^{q+1} - Y_2^q - Y_2, X_1^{q^2} - X_1, X_2^{q^2} - X_2\}$ is a Gröbner basis for $I_{q^2}^{(2)}$ with respect to $\prec_{w^{(2)}}$ giving us the basis

$$\{X_1^{i_1} Y_1^{j_1} X_2^{i_2} Y_2^{j_2} + I_{q^2} \mid 0 \leq i_1, i_2 < q^2, 0 \leq j_1, j_2 < q\}$$

for $\mathbb{F}_{q^2}[X_1, Y_1, X_2, Y_2]/I_{q^2}^{(2)}$ as a vectorspace over $\mathbb{F}_{q^2}$. For the code construction we need the following bijective evaluation map

$$\mathrm{EV} : \mathbb{F}_{q^2}[X_1, Y_1, X_2, Y_2]/I^{(2)} \to \mathbb{F}_{q^2}^{q^6}$$

given by $\mathrm{EV}(F(X_1, Y_1, X_2, Y_2) + I_{q^2}^{(2)}) = (F(Q_1,), \ldots, F(Q_{q^6}))$. Define $\Lambda^{(2)} = \Lambda \times \Lambda$ and $\left(\Lambda^{(2)}\right)^* = \Lambda^* \times \Lambda^*$. We have

$$\left(\Lambda^{(2)}\right)^* = w^{(2)}(\Delta_{\prec_{w^{(2)}}}(I_{q^2}^{(2)}))$$

where no two monomials in $\Delta_{\prec_{w^{(2)}}}(I_{q^2}^{(2)})$ have the same weight. Similar to the situation of a Hermitian code we consider a codeword $\boldsymbol{c} =$

$\mathrm{EV}(F(X_1,Y_1,X_2,Y_2)+I^{(2)}_{q^2})$ where without loss of generality we will assume that $F(X_1,Y_1,X_2,Y_2)\in \Delta_{\prec_{w^{(2)}}}(I^{(2)}_{q^2})$. We write $\lambda^{(2)}=(\lambda_1,\lambda_2)=w^{(2)}(\mathrm{lm}(F))$. We can estimate

$$\begin{aligned}|\Delta_{\prec_{w^{(2)}}}(\langle F(X_1,Y_1,X_2,Y_2)\rangle+I^{(2)}_{q^2})| &\leq |\Lambda^{(2)}-(\lambda^{(2)}+\Lambda^{(2)})|\\ &\leq q^6-(q^3-\lambda_1)(q^3-\lambda_2).\end{aligned}$$

Hence, $w_H(\boldsymbol{c})\geq (q^3-\lambda_1)(q^3-\lambda_2)$.
Consider the code $\widetilde{E}(\delta)$ which is to Hermitian codes what Massey-Costello-Justesen codes [13] are to Reed-Solomon codes

$$\begin{aligned}\widetilde{E}(\delta):=\ \mathrm{Span}_{\mathbb{F}_{q^2}}\Big\{&\mathrm{EV}(X_1^{i_1}Y_1^{j_1}X_2^{i_2}Y_2^{j_2}+I^{(2)}_{q^2})\mid 0\leq i_1,i_2<q^2, 0\leq j_1,j_2<q,\\ &(q^3-w(X_1^{i_1}Y_1^{j_1}))(q^3-w(X_2^{i_2}Y_2^{j_2}))\geq\delta\Big\}.\end{aligned}$$

From our discussion we conclude that the minimum distance satisfies $d(\widetilde{E}(\delta))\geq\delta$. To estimate the dimension we make use of the characterization (7). The task is to estimate the number of (λ_1,λ_2)s that satisfies $(q^3-\lambda_1)(q^3-\lambda_2)\geq\delta$. For this purpose we can replace Λ^* with

$$\{g,g+1,\ldots,q^3-1\}\cup\{\lambda_{n-g+1},\ldots,\lambda_n\}.$$

When estimating the dimension $k(\widetilde{E}(\delta))$ we shall furthermore ignore the elements in $\{\lambda_{n-g+1},\ldots,\lambda_n\}$. Writing $T=q^3-g$ we thereby get

$$\begin{aligned}k(\widetilde{E}(\delta)) &\geq |\{(i,j)\mid 0\leq i,j\leq T-1,(T-i)(T-j)\geq\delta\}|\\ &\geq \int_0^{T-\frac{\delta}{T}}\int_0^{T-\frac{\delta}{T-i}} dj di = T^2-\delta+\ln\left(\frac{\delta}{T^2}\right),\end{aligned}$$

where the last inequality holds under the assumption $\delta\geq T$.

Proposition 1. *Assume $\delta\geq T$ where $T=q^3-g$. The parameters of $\widetilde{E}(\delta)$ are $[n=q^6,k\geq T^2-\delta+\delta\ln(\delta/T^2),d\geq\delta]$.*

In [8] Feng-Rao improved codes $\widetilde{C}(\delta)$ over $\mathbb{F}_{q^2}[X_1,Y_1,X_2,Y_2]/I^{(2)}_{q^2}$ were considered and a formula similar to the above proposition was derived under a stronger assumption on δ. Feng-Rao improved codes are described by means of their parity check matrix which is not very useful when the aim is to construct a small-bias space. This is why we included the description of $\widetilde{E}(\delta)$ in the present paper. We have a proof that $\widetilde{E}(\delta)=$

$\widetilde{C}(\delta)$, however, we do not include it here as it has no implication for the construction of small-bias spaces. Observe that to derive Proposition 1 we did not use detailed information about the Weierstrass semigroup Λ but relied only on the genus and the number of roots of the Hermitian polynomial. Proposition 1 can be generalized to hold for not only two copies of Hermitian function fields but to arbitrary many such copies. Such constructions, however, are not useful when dealing with small-bias spaces so we do not treat them here.

From Proposition 1 and Theorem 2 we get a new class of ϵ-bias spaces:

Theorem 5. *For any ϵ, $0 < \epsilon < 1$ using codes $\widetilde{E}(\delta)$ as outer code in the construction of Theorem 2 one can construct ϵ-bias spaces with*

$$\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}, \qquad |\mathcal{X}| = \mathcal{O}\left(\left(\frac{k}{\epsilon + (1-\epsilon)\ln(1-\epsilon)}\right)^{\frac{4}{3}}\right). \tag{8}$$

Proof. In the following we will use the substitution $1 - \epsilon = \delta/N$ which follows from $\epsilon = (N - \delta)/N$. Assume $\delta > \sqrt{N}$. We then have $\delta > T$ which is the condition in Proposition 1. Note that $\delta > \sqrt{N}$ is equivalent to $\epsilon < 1 - (1/\sqrt{N})$. For $N \to \infty$ this becomes $\epsilon < 1$ which is actually no restriction at all. From the proposition we get

$$\begin{aligned}\frac{K}{N} &\geq \left(\frac{q^3 - g}{q^6}\right)^2 - \frac{\delta}{q^6} + \frac{\delta}{q^6}\ln\left(\frac{\delta}{(q^3-g)^2}\right)\\ &\geq o(1) + 1 - (1-\epsilon) + (1-\epsilon)\ln(1-\epsilon)\\ &= o(1) + \epsilon + (1-\epsilon)\ln(1-\epsilon).\end{aligned}$$

With $q^2 = 2^s$ we have

$$|\mathcal{X}| \leq \frac{2^s}{s}\left(\frac{k}{o(1) + \epsilon + (1-\epsilon)\ln(1-\epsilon)}\right).$$

But $|\mathcal{X}| = (2^s)^4$ implies $2^s = |\mathcal{X}|^{1/4}$ and (8) has been demonstrated.

Theorem 6. *Consider the family of ϵ-bias spaces in Theorem 5. Given $\alpha \in \mathbb{R}^+$ choose $\epsilon = k^{-\alpha}$ and let $k \to \infty$. We have*

$$\log_k(|\mathcal{X}|) = \frac{4}{3} + \frac{8}{3}\alpha + o(1). \tag{9}$$

Proof. We have

$$\log_k(|\mathcal{X}|) \leq \frac{4}{3} - \frac{4}{3}\log_k(\epsilon + (1-\epsilon)\ln(1-\epsilon)).$$

We now apply Taylors formula to derive $\ln(1-\epsilon) = -\epsilon - \epsilon^2/2(1-c)^2$ for some $c \in [0, \epsilon]$. This produces

$$\begin{aligned}\log_k(|\mathcal{X}|) &\leq \frac{4}{3} - \frac{4}{3}\log_k\left(\epsilon + (1-\epsilon)(-\epsilon - \frac{\epsilon^2}{2(1-\epsilon)^2})\right)\\ &\leq \frac{4}{3} - \frac{4}{3}\log_k\left(\epsilon^2\left(\frac{2(1-\epsilon)^2 - \epsilon^2}{(1-\epsilon)^2}\right)\right).\end{aligned}$$

With $\epsilon = k^{-\alpha}$ we arrive at (9).

4 Time complexity considerations

To build the multiset $\mathcal{X}$ in our construction we need to construct a generator matrix for the concatenated code. This involves the following tasks:

1. Build the generator matrix G_1 for $\widetilde{E}(\delta)$.
2. Express every entry of G_1 as a binary vector giving us G_2 (a matrix with binary vectors as entries).
3. For every row in G_2 we produce $s = \log_2(q^2)$ rows. This is done by taking cyclic shifts of all the vectors appearing in the row. We arrive at a matrix G_3.
4. Every entry in G_3 is a vector of length s and it must be multiplied with the $s \times 2^s$ generator matrix of the Walsh-Hadamard code producing G_4.

The total cost in binary operations is estimated as follows:

1. Determining functions and points for the code construction is inexpensive. To produce one entry costs $\mathcal{O}\left(\log(N)\log(\log(N))\right)$ operations. G_1 is a $K \times N$ matrix. Using $K \leq N - D + 1$, $\epsilon = (N-D)/N$, $\epsilon = k^{-\alpha}$, and $k = K\log_2(N)/6$ we arrive at $K \leq N^{\frac{1}{1+\alpha}}(\log_2(N))^{\frac{-\alpha}{1+\alpha}}6^{\frac{\alpha}{1+\alpha}}$. So the price for building G_1 is $\mathcal{O}\left(N^{\frac{2+\alpha}{1+\alpha}}(\log(N))^{\frac{1}{1+\alpha}}\log(\log(N))\right)$.
2. To produce one entry in G_2 costs $\mathcal{O}\left(N^{\frac{1}{3}}\log(N^{\frac{1}{3}})\log(\log(N^{\frac{1}{3}}))\right)$ operations. That is, to produce G_2 from G_1 amounts to $\mathcal{O}\left(N^{\frac{7+4\alpha}{3+3\alpha}}\log(N)^{\frac{1}{1+\alpha}}\log(\log(N))\right)$ operations.
3. There will be $\mathcal{O}\left(N^{\frac{2+\alpha}{1+\alpha}}(\log(N))^{\frac{1}{1+\alpha}}\right)$ entries in G_3 each coming with a cost of s operations. Altogether we have $\mathcal{O}\left(N^{\frac{2+\alpha}{1+\alpha}}(\log(N))^{\frac{2+\alpha}{1+\alpha}}\right)$ operations.

4. The price for multiplying with a generator matrix for the Walsh-Hadamard code is $N^{\frac{1}{3}} \log(N)$ giving a total cost of
$$\mathcal{O}\left(N^{\frac{7+4\alpha}{3+3\alpha}} (\log(N))^{\frac{2+\alpha}{1+\alpha}}\right) \tag{10}$$
operations for producing G_4 from G_3.

Clearly, the overall cost is that of (10). Note that (10) counts binary operations in contrast to (6) which counts operations in $\mathbb{F}_{q^2}$.

5 Small-bias spaces from norm-trace codes

The method developed by Ben-Aroya and Ta-Shma for Hermitian codes in [2] were generalized to norm-trace codes by Matthews and Peachey in [12]. Given $r \geq 2$ consider the C_{ab} curve [11]
$$X^{\frac{q^r-1}{q-1}} - Y^{q^{r-1}} - Y^{q^{r-2}} - \cdots - Y^q - Y$$
known as the norm-trace curve over $\mathbb{F}_{q^r}$ [6]. Clearly, $r = 2$ corresponds to the Hermitian function field. The following theorem from [12] coincides with (3) when $l = 4$.

Theorem 7. *Given an integer l, $l \geq 4$, define $r = \lfloor (l+2)/3 \rfloor$. Let k be a positive integer and ϵ a real number, $0 < \epsilon < 1$ such that*
$$\frac{\epsilon}{\left(\log_v(1/\epsilon)\right)^{\frac{1}{\sqrt{l}}}} \leq k^{\frac{-1}{\sqrt{l}}} \tag{11}$$
holds. Here, v is any fixed real number larger than 1. Using the norm-trace function field over $\mathbb{F}_{q^r}$ one can construct an ϵ-bias space $\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}$ with
$$|\mathcal{X}| = \mathcal{O}\left(\left(\frac{k}{\epsilon^{l-\sqrt{l}} \log_v(1/\epsilon)}\right)^{\frac{l+1}{l}}\right).$$

In the above theorem it is not completely clear how well the cases $l \geq 5$ compete with the case $l = 4$. Below we address this question and also compare the small-bias spaces from Theorem 7 with those achieved by using the codes $\widetilde{E}(\delta)$ as is done in the present paper.

We first translate Theorem 7 into the setting from Section 1 where for increasing k and fixed α we consider a sequence of ϵ-bias multisets with $\epsilon = k^{-\alpha}$. Condition (11) from Theorem 7 then translates into
$$k^{1-\alpha\sqrt{l}} \leq \alpha \log_v(k).$$

For fixed v, $\log_v(k) = \mathcal{O}\left(k^{\beta}\right)$ holds for any $\beta > 0$. Therefore we have

$$1 - \alpha\sqrt{l} \leq \log_k(\alpha).$$

Letting $k \to \infty$ we get the condition

$$\frac{1}{\sqrt{l}} \leq \alpha.$$

Theorem 7 therefore guarantees that for any $\alpha \geq 1/\sqrt{l}$ we can construct an infinite sequence of ϵ-bias spaces with $\epsilon = k^{-\alpha}$, $\mathcal{X} \subseteq \mathbb{F}_2^{\Omega(k)}$ such that

$$\log_k(|\mathcal{X}|) = \frac{l+1}{l}(1 + \alpha(l - \sqrt{l})) + o(1). \tag{12}$$

Given an α and two integers $l_1, l_2 \geq 4$ with $\alpha \geq 1/\sqrt{l_i}$, $i = 1, 2$ it is clear from (12) that the best result is obtained by choosing the smallest l_i. So the advantage of Theorem 7 over (3) boils down to the fact that Theorem 7 allows for any α provided that the l is chosen accordingly while (3) requires $\alpha \geq 1/2$. Recall from Section 3 that using the code $\widetilde{E}(\delta)$ in the construction of Theorem 2 one achieves

$$\log_k(|\mathcal{X}|) = \frac{4}{3} + \frac{8}{3}\alpha + o(1) \tag{13}$$

for any choice of α. We now compare this result with (12) ignoring of course the $o(1)$ parts. For fixed l (12) is a linear expression in α which is smaller than the linear expression from (13) when $\alpha = 0$. We now show that for $\alpha = 1/\sqrt{l}$ (which is the smallest α allowed) (12) is larger than (13) when $l \geq 5$. It follows that none of the cases $l \geq 5$ can compete with the construction of the present paper. To show that (12) is larger than (13) for $\alpha = 1/\sqrt{l}$ we substitute $k = \sqrt{l}$ into (12)-(13) to get

$$\frac{1}{k^2}(k^3 - \frac{4}{3}k^2 - \frac{5}{3}k).$$

The function $k^3 - \frac{4}{3}k^2 - \frac{5}{3}k$ is positive for k belonging to the interval from 0 to approximately 2.119 and negative for higher values of k. Therefore for all $l \geq 5$ indeed (13) is better than (12).

6 Acknowledgments

The present work was done while Ryutaroh Matsumoto was visiting Aalborg University as a Velux Visiting Professor supported by the Villum Foundation. The authors gratefully acknowledge this support. The authors also gratefully acknowledge the support from the Danish National Research Foundation and the National Science Foundation of China (Grant No. 11061130539) for the Danish-Chinese Center for Applications of Algebraic Geometry in Coding Theory and Cryptography.

References

1. N. Alon, O. Goldreich, J. Hastad, and R. Peralta: Simple constructions of almost k-wise independent random variables. *Random Structures Algorithms* **3** (1992), no. 3, 289-303.
2. A. Ben-Aroya and A. Ta-Shma: Constructing small-bias sets from algebraic- geometric codes. *FOCS'2009*, 191-197.
3. D. Cox, J. Little and D. O'Shea: *Ideals, Varieties, and Algorithms, Sec. Ed.*, Springer, 1997.
4. J. Fitzgerald and R. F. Lax: Decoding Affine Variety Codes Using Gröbner Bases. *Des. Codes Cryptography,* **13**, 1998, 147-158.
5. A. Garcia and H. Stichtenoth: On the asymptotic behaviour of some towers of function fields over finite fields. *J. Number Theory,* **61**, 1996, 248-273.
6. O. Geil: On codes from norm-trace curves. *Finite Fields and their Applications* **9** (2003), 351-371.
7. O. Geil and T. Høholdt: Footprints or Generalized Bezout's Theorem. *IEEE Trans. Inform. Theory,* **46**, 2000, 635-641.
8. O. Geil and T. Høholdt: On Hyperbolic Type Codes. *Proceedings of 2003 IEEE International Symposium on Inf. Theory,* Yokohama, 2003, 331.
9. T. Høholdt: On (or in) Dick Blahut's 'footprint', in "Codes, Curves and Signals," (A. Vardy, Ed.), Kluwer Academic, Norwell, MA, 1998, 3-9.
10. T. Høholdt, J. van Lint and R. Pellikaan: Algebraic Geometry Codes, Chapter 10 in "Handbook of Coding Theory," (V.S. Pless and W.C. Huffman, Eds.), vol. 1, Elsevier, Amsterdam, 1998, 871-961.
11. S. Miura and N. Kamiya: Geometric-Goppa codes on some maximal curves and their minimumdistance. *Proc. of 1993 IEEE Inf. Th. Workshop* Susonon-shi, Shizuoka, Japan, June 4-8, 1993, 85-86.
12. G. L. Matthews and J. Peachey: Small-bias sets from extended norm-trace codes. To appear in Proceedings of Fq10, *Contemporary Mathematics*, AMS.
13. J. Massey, D. J. Costello, and J. Justesen: Polynomial Weights and Code Constructions. *IEEE Trans. Inf. Theory*, **19**, 1973, 101-110.
14. R. Meka and D. Zuckerman: Small-Bias Spaces for Group Products. *APPROX-RANDOM* 2009, 658-672.
15. J. Naor and M. Naor: Small-bias probability spaces: eficient construction and applications. *SIAM J. Comput.* **22** (1993), 838-856.
16. K. W. Shum, I. Aleshnikov, P. Vijay Kumar, H. Stichtenoth, and V. Deolalikar: A Low-Complexity Algorithm for the Construction of Algebraic-Geometric Codes Better Than the Gilbert-Varshamov Bound. *IEEE Trans. Inform. Theory,* **47**, 2001, 2225-2241.
17. U. V. Vazirani: Randomness, adversaries, and computation, Ph.D. thesis, EECS, UC Berkeley, 1986.